ランダムウォークの座標の推定

樋口さぶろお

龍谷大学理工学部数理情報学科

計算科学☆実習 B L03(2021-04-22 Thu)

最終更新: Time-stamp: "2021-06-28 Mon 16:27 JST hig"

今日の目標

- ランダムウォークの座標の母平均値, 母分散, 母期待値, 母比率を C と Excel で推定できる
- 母平均値, 母分散は正確に計算できる

L02-Q2
Quiz 解答:ランダムウォーカーの到達点の座標の母平均・母分散

1. 標本平均値 $\overline{X(3)} = \frac{1}{10}(3+3+\cdots+(-3)) = 1$. よって, 母平均値 $\mathrm{E}[X(3)]$ は 1 と推定できる.
2. 不偏標本分散 $S^2 = \frac{1}{10-1}((3-1)^2+\cdots+(-3-1)^2) = \frac{32}{9}$. よって母分散 $\mathrm{E}[X(3)]$ は $\frac{32}{9}$ と推定できる.
3. 標本期待値 $\overline{X(3)^3} = \frac{1}{10}(3^3+\cdots+(-3)^3) = \frac{29}{5}$. よって母期待値 $\mathrm{E}[X(3)^3]$ は $\frac{29}{5}$ と推定できる.
4. 標本期待値 $\overline{\mathrm{I}_{[X>1]}(X(3))} = \frac{1}{10}(1+1+1+0+\cdots+0) = \frac{3}{10}$. よって母比率 $p = \mathrm{E}[\mathrm{I}_{[X>1]}(X(3))]$ は $\frac{3}{10}$ と推定できる.

ここまで来たよ

3 ランダムウォークのサンプルパスの標本抽出

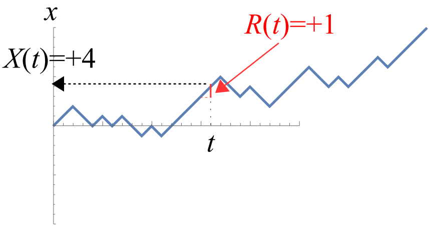
3 ランダムウォークの座標の推定
- 座標の推定
- ランダムウォークの座標の母平均値と母分散
- ソリューションとプロジェクト
- 乱数のシードとは?

ランダムウォーク (確率過程の例)

ランダムウォーク

$$X(t) = X(t-1) + R(t), \quad X(0) = x_0$$

階差数列 $R(t)$: 独立同分布にしたがう確率変数.
ランダムウォーカー座標 $X(t)$: 確率変数.

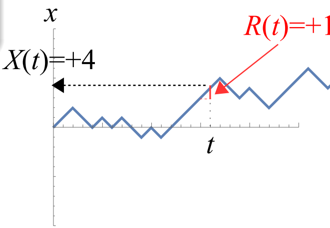


ランダムウォークの座標の母期待値，母比率は?

$R(t)$ やランダムウォークの座標 $X(1000)$ は確率変数.

母ナントカの問

母ナントカを厳密に求めよう.



- 母平均値 $\mathrm{E}[R(t)]$, 母期待値 $\mathrm{E}[\mathrm{e}^{R(t)}]$, 母比率 $P(R(t) > 1)$
- 母平均値 $\mathrm{E}[X(1000)]$, 母期待値 $\mathrm{E}[\mathrm{e}^{X(1000)}]$, 母比率 $P(X(1000) > 1)$
- 母比率 $P(X(50) = 12$ かつ $X(100) = 25)$

別のタイプの問 (推定)

$R(t)$ の確率分布 $f(x)$ の式を知らない (例:誰かが作った getrandom の中身不明) けど, Excel データ (標本) だけはあるとする. 母ナントカを推定したい.

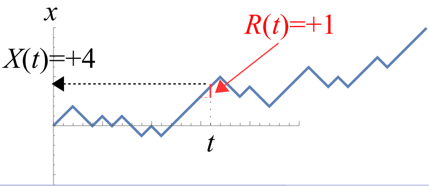
ランダムウォークの座標 $X(t)$ の標本

$X(t)^{(n)}$

t: 時刻, 数列の第 t 項

(n): サンプル内通し番号

	$t=0$	$t=1$	$\cdots$	$t=T$
$n=1$	$X(0)^{(1)}$,	$X(1)^{(1)}$,	$\cdots$	$X(T)^{(1)}$, 改行
$n=2$	$X(0)^{(2)}$,	$X(1)^{(2)}$,	$\cdots$	$X(T)^{(2)}$, 改行
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$n=N$	$X(0)^{(N)}$,	$X(1)^{(N)}$,	$\cdots$	$X(T)^{(N)}$, 改行



Excel で $X(t)$ の母平均値，母分散，母比率を推定

確率統計☆演習 I(2020)L02

- 標本平均値，標本期待値 `average`
- 不偏標本分散 `var.s`
- $\mathrm{I}_{[\text{条件}]}(x) = \begin{cases} 1 & (x\ \text{が条件を満たす}) \\ 0 & (x\ \text{が条件を満たさない}) \end{cases}$
  `=if(x の条件,1,0)`

$X(0)$ と $X(1)$ は別の確率変数だから，「混ぜて標本ナントカを得る」ことに意味はない．

標本期待値は Excel を使わなくても C の中でも計算できる!

$$\overline{\phi(X(T))} = \frac{1}{N}\sum_{n=1}^{N} \phi(X(T)^{(n)})$$

```
                                  /*1*/
for(n){
                                  /*2*/
  for(t){
                                  /*3*/
    x=x+getrandom(getuniform());
                                  /*4*/
  }
                                  /*5*/
}
                                  /*6*/
```

sum1=0, sum1+=phi(x) (例 $\phi(x) = x^3$), printf("%f",(double)sum1/N)?

標本比率の C による計算

$\phi(x) = \mathrm{I}_{[条件]}(x)$ に相当するに相当する `phi` は?
例で. $P(X > 10)$ を考えるとき,

$$\mathrm{I}_{[x>10]}(x) = \begin{cases} 1 & (x > 10) \\ 0 & (x \geq 10) \end{cases}$$

↓

```c
int phi(int x){
  if(x>10){
    return 1;
  } else {
    return 0;
}
```

そのままじゃん.
この場合, sum1 は count1 と名付けたほうがいいかも.

ここまで来たよ

L03-Q1

Quiz(ランダムウォーカーの到達点の座標の母平均・母分散)

時刻 t におけるランダムウォーカーの座標 $X(t)$ を, 次の漸化式で定める ($t = 0, 1, 2, \dots$).

$$X(t) = X(t-1) + R(t), \quad X(0) = 0$$

確率変数 $R(t)$ ($t = 1, 2, 3, \dots$) は, 互いに独立, 同分布に従い,

- 確率 $5/9$ で $R(t) = -1$,
- 確率 $1/9$ で $R(t) = 0$,
- 確率 $3/9$ で $R(t) = +1$,

の値をとる.

1. $R(t)$ の母平均値を求めよう.
2. $R(t)$ の母分散を求めよう.
3. $R(t)$ の母標準偏差を求めよう.
4. $X(20)$ の母平均値を求めよう.
5. $X(20)$ の母分散を求めよう.
6. $X(20)$ の母標準偏差を求めよう.

離散型確率変数の母期待値 岩薩林 確率・統計 (3.3) 母分散 岩薩林 確率・統計 (3.7),(3.8) 独立同分布の和の分散 岩薩林 確率・統計 (3.21)

ランダムウォーク

ランダムウォークの定義

$R(t)$: 独立同分布に従う離散型確率変数. $t = 1, 2, 3, \dots$
$X(t)$: 次で決まる確率変数.

初期条件 $X(a) = b$　(正確には $P(X(a) = b) = 1$)
漸化式 $X(t) = X(t-1) + R(t)$　$(t = a+1, a+2, a+3, \dots)$

$R(t)$ が 二項分布 岩薩林 確率・統計 §3.4 $\mathrm{B}(1, p)$ に
したがうとき

$$\text{確率 } P(R(t) = r) = \begin{cases} p & (r = 1) \\ q = 1 - p & (r = 0) \\ 0 & (\text{他}) \end{cases}$$

例 $p = \frac{2}{3}$.

ランダムウォークの座標の母平均値と母分散

確率統計☆演習 I(2020)L06

$$
\begin{aligned}
X(t) =& X(t-1) + R(t) \\
=& (X(t-2) + R(t-1)) + R(t) \\
=& \cdots = X(a) + R(a+1) + \cdots + R(t)
\end{aligned}
$$

連続/離散型ランダムウォークの母平均値/母分散

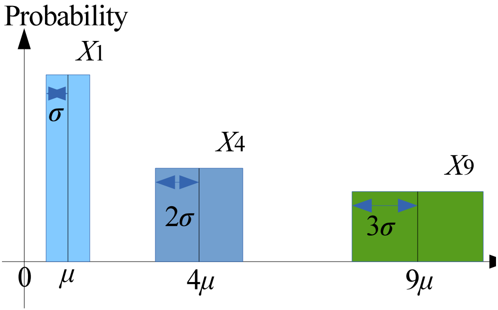
$\mathrm{E}[R(t)] = \mu, \mathrm{V}[R(t)] = \sigma^2$ のとき,

$$
\begin{aligned}
\mathrm{E}[X(T)] =& \mathrm{E}[X(a) + R(a+1) + \cdots + R(T)] = b + (T-a)\mu \quad (\mathrm{E}[] \text{ の性質}), \\
\mathrm{V}[X(T)] =& \mathrm{V}[X(a) + R(a+1) + \cdots + R(T)] \\
=& \mathrm{V}[R(a+1) + \cdots + R(T)] \quad (X(a) = b \text{ は定数}) \\
=& (T-a)\sigma^2 \quad (R(t) \text{ は互いに独立})
\end{aligned}
$$

$a = b = 0$ という簡単な場合, $X(t)$ の母平均値, 母分散は t に正比例する.

$X(t)$ の母標準偏差 $= \sqrt{\mathrm{V}[X(T)]} =$ 自分で書いてね

ってことは，ランダムウォークの座標の確率分布の時間変化はこんな感じ？



L03-Q2

Quiz(ランダムウォーカーの到達点の座標の母平均・母分散)

時刻 $t=-5$ に, $x=2$ から出発するランダムウォークを考える.
時刻ごとに, ランダムウォーカーは,

- 確率 $2/3$ で $+1$,
- 確率 $1/3$ で 0,

だけ移動する. 時刻ごとの移動分は独立である.

1. 確率 $P(X(3)=5)$ を求めよう.
2. $X(3)$ の母平均値を求めよう.
3. $X(3)$ の母分散を求めよう.
4. $X(3)$ の母標準偏差を求めよう.

ここまで来たよ

3 ランダムウォークのサンプルパスの標本抽出

3 ランダムウォークの座標の推定
- 座標の推定
- ランダムウォークの座標の母平均値と母分散
- ソリューションとプロジェクト
- 乱数のシードとは?

IDE, ソリューション, プロジェクト

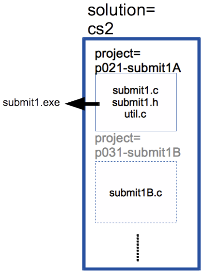
IDE=Integrated Develpment Environment 統合開発環境

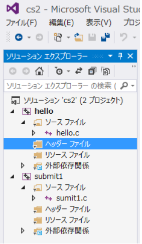
編集, コンパイル, 実行, デバッグを 1 つのアプリケーション内で完結して実行できるようにしたもの．例: Visual Studio, Eclipse, NetBeans.

プロジェクト=1 個の実行ファイル (a.out, *.exe) を作るための，複数のソースファイル (*.c, *.h) をまとめて管理するもの．プロジェクト内に main は 1 個だけ．

ソリューション=複数の関係するプロジェクトをまとめて管理するもの

VS では, 1 個のソリューションを開いた状態で，スタートアッププロジェクトをコンパイル, 実行する.





ここまで来たよ

3 ランダムウォークのサンプルパスの標本抽出

3 ランダムウォークの座標の推定
- 座標の推定
- ランダムウォークの座標の母平均値と母分散
- ソリューションとプロジェクト
- 乱数のシードとは?

擬似乱数列生成の仕組みとシード

擬似乱数列=‘ほぼ’ ランダムな数列

2 回続けて `int rand()` を呼んで得られる 2 個の数の分布は独立であるかのように考えてる (けど… そこが擬似).

(標本抽出のたびに別々の)seed を指定することが必要

- seed に応じて, 毎回異なる乱数列が得られる.
- 特定の乱数列に対する動作を再現できる. デバッグでは必須.
- seed を適切に設定すると, 複数回の実行で, 別々の (独立な) 標本抽出が行える.

L03-Q3

Quiz(rand() の振る舞い)

次のプログラムで, seed を無作為に入力するとき, A が出力される確率は?

```
int getrandom(double y){
  if( y<1/3.0 ){
    return 0;
  }else{
    return 1;
  }
}

int main(){
  int seed;
  scanf("%d",&seed);
  srand(seed);
  if( getrandom(getuniform())==getrandom(getuniform()) ){
    printf("A\n");
  }
  return 0;
}
```

1. 0
2. 1/2
3. 5/9 に近い
4. 6/9 くらい
5. 1

Moodle App

取得 `https://download.moodle.org/mobile`

指定する URL `https://moodle.hig3.net/moodle`
Moodle から通知がくる, 閲覧しやすい.
テストや練習問題には使わないほうがいい?